

# A Linear Programming Approach to Dynamical Equivalence, Linear Conjugacy, and the Deficiency One Theorem

Matthew D. Johnston  
 Department of Mathematics  
 San José State University  
 One Washington Square  
 San José, CA 95192  
 matthew.johnston@sjsu.edu

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Background</b>	<b>3</b>
2.1	Chemical Reaction Networks . . . . .	3
2.2	Mass Action Systems . . . . .	4
2.3	Dynamical Equivalence and Linear Conjugacy . . . . .	5
2.4	Deficiency Theory . . . . .	6
<b>3</b>	<b>Main Results</b>	<b>8</b>
3.1	Mixed-Integer Linear Programming Framework . . . . .	9
3.2	Initialization of Program . . . . .	9
3.3	Implementing Dynamical Equivalence and Linear Conjugacy . . . . .	10
3.4	Implementing Deficiency Conditions . . . . .	10
3.5	Implementing One Terminal Strong Linkage Class . . . . .	13
<b>4</b>	<b>Examples</b>	<b>14</b>
<b>5</b>	<b>Conclusions</b>	<b>17</b>

## Abstract

The well-known Deficiency One Theorem gives structural conditions on a chemical reaction network under which, for any set of parameter values, the steady states of the corresponding mass action system may be easily characterized. It is also known, however, that mass action systems are not uniquely associated with reaction networks and that some representations may satisfy the Deficiency One Theorem while others may not. In this paper we present a mixed-integer linear programming framework capable

of determining whether a given mass action system has a dynamically equivalent or linearly conjugate representation which has an underlying network satisfying the Deficiency One Theorem. This extends recent computational work determining linearly conjugate systems which are weakly reversible and have a deficiency of zero.

**Keywords:** chemical reaction networks; chemical kinetics; deficiency; linear programming; dynamical equivalence

**AMS Subject Classifications:** 80A30, 90C35

## 1 Introduction

A chemical reaction network is given by sets of reactants which interact according to fixed reaction channels to form new sets of reactants. Under suitable kinetic assumptions, such as spatial homogeneity and sufficient molecularity, these networks can be modeled by a system of autonomous polynomial ordinary differential equations known as a mass action system. The study of mass action systems, and the related area of chemical reaction network theory (CRNT), has been studied increasingly in recent years as the interdisciplinary area of systems biology has become more prominent.

A primary focus of CRNT is on the relationship between the topological structure of the network of interactions and the permissible dynamical behaviors of the corresponding reaction systems under a variety of kinetic assumptions. The canonical paper [18], published in 1972, introduces the notion of complex-balancing in a reaction network and shows that this condition is sufficient to guarantee very strong asymptotic behavior of the corresponding mass action system. The concurrent papers [8] and [17] introduce a nonnegative network parameter known as the deficiency and develop the now classical *Deficiency Zero Theorem*. This theorem states that a network which is weakly reversible and has a deficiency of zero is complex balanced for all parameter values, and therefore the corresponding mass action systems possess the associated asymptotic behavior. Networks with a higher deficiency, and in particular a deficiency of one, have also been studied. The most well-known of these is the *Deficiency One Theorem*, which gives conditions sufficient for the uniqueness of steady states of higher deficiency mass action systems [9, 11]. Other deficiency-based results characterizing the existence, number, and nature of steady states have also been derived [4, 10, 19, 29].

It is furthermore known that two chemical reaction networks can generate the same mass action system and therefore be dynamically equivalent. In such cases, one network may have a desirable network property such as weak reversibility or a low deficiency while another might not. Significant work has consequently been conducted on determining network representations of mass action systems with desirable structural properties. Mixed-integer linear programming (MILP) frameworks are now established for determining dynamically equivalent network structures which are linearly conjugate [24], weakly reversible [2, 24, 28, 34], detailed and complex balanced [23, 32], reaction dense and reaction sparse [33], and have a minimal deficiency [22, 25]. In particular, the question of whether a mass action system has a network representation satisfying the assumptions of the Deficiency Zero Theorem was answered in [22]. Many of these features are implemented in the computational package CRNreals [31].

One question which remains open is whether a given a mass action system has a dynamically equivalent or linearly conjugate representation satisfying the Deficiency One Theorem. To date, there exists no framework by which to check the following technical assumptions of the theorem:

(a) that the sum of the deficiencies of each linkage classes is bounded by one and sums to the overall network deficiency; and (b) that each linkage class contains a single terminal strong linkage class. In this paper, we extend the MILP framework introduced in the papers outlined above to include conditions (a) and (b), and therefore determine whether a network has a representation satisfying the assumptions of the Deficiency One Theorem. We present examples of mass action systems and networks which, while not amenable to the Deficiency One Theorem directly, have dynamically equivalent and/or linearly conjugate systems which are amenable to it.

## 2 Background

In this section, we give the necessary terminology, notation, and background results relevant to the results contained in Section 3.

### 2.1 Chemical Reaction Networks

The following is the fundamental object of this paper.

**Definition 2.1.** A chemical reaction network is a triple of sets  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  where:

1. The **species set**  $\mathcal{S} = \{X_1, \dots, X_m\}$  consists of the elementary chemical species capable of undergoing chemical change.
2. The **complex set**  $\mathcal{C} = \{C_1, \dots, C_n\}$  consists of linear combinations of the species of the form

$$C_i = \sum_{j=1}^m y_{ij} X_j$$

where  $y_{ij} \geq 0$  are the **stoichiometric coefficients**. We will let  $y_i = (y_{i1}, \dots, y_{im}) \in \mathbb{R}_{\geq 0}^m$  denote the vector of stoichiometric coefficients corresponding to the complex  $C_i \in \mathcal{C}$ .

3. The **reaction set**  $\mathcal{R} = \{R_1, \dots, R_r\}$  consists of ordered pairs of complexes  $(C_i, C_j) \in \mathcal{R}$ . Reactions will also be represented as  $C_i \rightarrow C_j \in \mathcal{R}$ .

We will assume that every reaction network satisfies the following: (i) every species is contained in at least one complex; and (ii) every complex is contained in at least one reaction.

Note that we do not include the typical assumption that there are no self-reactions in the network (i.e. reactions of the form  $C_i \rightarrow C_i$ ). The algorithm presented in Section 3 will require that we permit complexes which are isolated from every other complex (i.e. they are not connected to any other complex by any reaction). We will accommodate these isolated complexes by associating each such complex to a self-reaction. Such an allowance does not limit our ability to apply the Deficiency One Theorem (see Lemma 3.1).

Associated with each chemical reaction network  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  is a digraph  $G(V, E)$  where  $V = \mathcal{C}$  and  $E = \mathcal{R}$ . This digraph is known in the literature as the reaction graph of the network [18]. Many properties from graph theory have been studied in the context of chemical reaction networks which we now briefly introduce.

Two complexes  $C_i$  and  $C_j$  are said to be connected if there is a sequence of complexes such that  $C_i = C_{\mu(1)} \leftrightarrow C_{\mu(2)} \leftrightarrow \dots \leftrightarrow C_{\mu(l)} = C_j$  where  $C_{\mu(k-1)} \leftrightarrow C_{\mu(k)}$  implies either  $C_{\mu(k-1)} \leftarrow C_{\mu(k)}$  or

$C_{\mu(k-1)} \rightarrow C_{\mu(k)}$ . We say there is a path from  $C_i$  to  $C_j$  if there is a sequence of complexes such that  $C_i = C_{\mu(1)} \rightarrow C_{\mu(2)} \rightarrow \dots \rightarrow C_{\mu(l)} = C_j$ . A subset of complexes  $L \subseteq \mathcal{C}$  is called a linkage class if it is maximally connected. The set of linkage class of a network will be denoted  $\mathcal{L} = \{L_1, \dots, L_\ell\}$ . A subset of complexes  $L \subseteq \mathcal{C}$  is called a strong linkage class if it is maximally path-connected; that is, for every two complexes  $C_i, C_j \in L$ ,  $C_i \neq C_j$ , a path from  $C_i$  to  $C_j$  implies a path from  $C_j$  to  $C_i$ , but  $C_k \notin L$  implies either no path from  $C_i$  to  $C_k$  or no path from  $C_k$  to  $C_i$ . A strong linkage class is called terminal if there is no reaction from a complex in the strong linkage class to a complex not in the strong linkage class. The set of terminal strong linkage classes will be denoted  $\mathcal{T} = \{T_1, \dots, T_t\}$ . A network is said to be weakly reversible if every linkage class is a strong linkage class.

To every reaction  $C_i \rightarrow C_j \in \mathcal{R}$  there is an associated reaction vector  $y_j - y_i \in \mathbb{R}^m$  which tracks the net gain/loss of each species as a result of a single instance of the reaction. The stoichiometric subspace of the network is given by:

$$S = \text{span}\{(y_j - y_i) \in \mathbb{R}^m \mid C_i \rightarrow C_j \in \mathcal{R}\}.$$

The dimension of the stoichiometric subspace will be denoted  $s = \dim(S)$ .

## 2.2 Mass Action Systems

In order to determine how the species involved in a chemical reaction network evolve over time, it is necessary to make assumptions on the kinetics of the system. It is common to assume that the system is subject to the law of mass action, which states that the rate of a reaction is proportional to the product of the reactant concentrations. For example, a reaction of the form  $X_1 + X_2 \rightarrow \dots$  would have the rate  $k \cdot x_1 x_2$  where  $x_1 = [X_1]$  and  $x_2 = [X_2]$  are the concentrations of  $X_1$  and  $X_2$ , respectively. Although mass action kinetics is the most common kinetic form, others are commonly used in biochemistry, including Michaelis-Menten kinetics [27] and Hill kinetics [16].

**Definition 2.2.** Consider a chemical reaction network  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ . Let  $\mathbf{x}(t) = (x_1(t), \dots, x_m(t)) \in \mathbb{R}_{\geq 0}^m$  denote the vector of reactant concentrations at time  $t \geq 0$ , and  $\mathcal{K} = \{k(i, j) > 0 \mid C_i \rightarrow C_j \in \mathcal{R}\}$  denote a set of rate constants. Then the **mass action system**  $(\mathcal{S}, \mathcal{C}, \mathcal{R}, \mathcal{K})$  associated with the network  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  and rate constant set  $\mathcal{K}$  is given by

$$\frac{d\mathbf{x}}{dt} = Y \cdot A(\mathcal{K}) \cdot \Psi(\mathbf{x}(t)) \quad (1)$$

where

1.  $Y \in \mathbb{Z}_{\geq 0}^{m \times n}$  is the **stoichiometric matrix** with entries  $Y_{i,j} = y_{ji}$ ;
2.  $A(\mathcal{K}) \in \mathbb{R}^{n \times n}$  is the **Laplacian matrix** with entries

$$[A(\mathcal{K})]_{i,j} = \begin{cases} -\sum_{l=1}^r k(i, l), & \text{if } i = j \\ k(j, i), & \text{otherwise;} \end{cases}$$

3.  $\Psi(\mathbf{x}) \in \mathbb{R}_{\geq 0}^n$  is the vector with entries  $\Psi_i(\mathbf{x}) = \prod_{j=1}^m x_j^{y_{ij}}$ .

The form (1) emphasizes the connectivity structure of network. In particular, we have  $[A(\mathcal{K})]_{ji} > 0$  for  $i \neq j$  if and only if  $C_i \rightarrow C_j \in \mathcal{R}$ . That is, the distribution of zero and non-zero off-diagonal elements of  $A(\mathcal{K})$  encode the structure of the network.

It can also be easily shown that (1) can be represented in the alternative, and somewhat more intuitive, form:

$$\frac{d\mathbf{x}}{dt} = \sum_{C_i \rightarrow C_j \in \mathcal{R}} k(i, j) (y_j - y_i) \prod_{l=1}^m x_l^{y_{il}}. \quad (2)$$

It follows immediately from (2) that  $\mathbf{x}'(t) \in S$ . Nonnegativity of solutions is also well-known so that  $\mathbf{x}(t) \in \mathbf{C}_{\mathbf{x}_0}$  for all trajectories  $\mathbf{x}(t)$  of (1), where  $\mathbf{C}_{\mathbf{x}_0} = (\mathbf{x}_0 + S) \cap \mathbb{R}_{>0}^m$  is the stoichiometric compatibility class associated with the initial condition  $\mathbf{x}_0 \in \mathbb{R}_{>0}^m$  [35].

### 2.3 Dynamical Equivalence and Linear Conjugacy

It is well-known that structurally distinct chemical reaction networks can generate the same mass action system (1) under the assumption of mass action kinetics [6, 18, 21]. For example, consider the following edge-weighted networks where the weights correspond to the value of the rate constant:

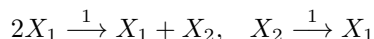


and

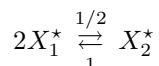


It can easily be seen that both of these networks generate the system of differential equations  $\dot{x}_1 = -\dot{x}_2 = -2x_1^2 + 2x_2^2$  under the assumption of mass action kinetics (1). The networks are therefore said to be dynamically equivalent. Note, however, that the connectivity properties of (3) and (4) are different. In particular, (4) is weakly reversible while (3) is not.

The notion of linear conjugacy of mass action systems was introduced in [21]. Two systems  $(\mathcal{S}, \mathcal{C}, \mathcal{R}, \mathcal{K})$  and  $(\mathcal{S}^*, \mathcal{C}^*, \mathcal{R}^*, \mathcal{K}^*)$  are said to be linearly conjugate to one another if the trajectories  $\mathbf{x}(t)$  and  $\mathbf{x}^*(t)$  of (1) are related by  $x_i(t) = c_i x_i^*(t)$ ,  $i = 1, \dots, n$ , for some constants  $c_i > 0$ . For example, the networks



and



have the non-dynamically equivalent mass action systems  $\dot{x}_1 = -\dot{x}_2 = -x_1^2 + x_2$  and  $\dot{x}_1^* = -2\dot{x}_2^* = -(x_1^*)^2 + 2x_2^*$ , respectively. These systems are linearly conjugate to one another under the transformation  $x_1(t) = x_1^*(t)$  and  $x_2(t) = 2x_2^*(t)$ . Linearly conjugate systems share many qualitative properties, including the number and stability of positive steady states, and the properties of persistence and boundedness (see Lemma 3.2 of [21]).

Since dynamically equivalent systems are a subset of linearly conjugate systems taking  $c_i = 1$  for all  $i = 1, \dots, m$ , in the rest of this paper we will only refer to the study of linearly conjugate systems. We will do this with the understanding that the systems studied may, in fact, be dynamically equivalent.

## 2.4 Deficiency Theory

The following network parameter was introduced in [8] and [17], and has been studied significantly since [5, 9, 11, 29].

**Definition 2.3.** *The deficiency of a chemical reaction network  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  is given by  $\delta = n - \ell - s$  where  $n = |\mathcal{C}|$ ,  $\ell = |\mathcal{L}|$ , and  $s = \dim(S)$ .*

The deficiency is a nonnegative integer which may be determined based on study of the network topology alone. That is, it is independent of the rate constants and even the assumption of mass action kinetics (e.g. Michaelis-Menten, Hill kinetics). Further connections between the deficiency, different rate forms, and the matrices  $Y$  and  $A_k$  from Definition 2.2 are well-known but will not be summarized here [3, 8, 14].

The following classical result was first presented in [8, 17, 18].

**Theorem 2.1** (Deficiency Zero Theorem). *Consider a chemical reaction network  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  which satisfies the following:*

1. *the network is weakly reversible; and*
2. *the deficiency is zero (i.e.  $\delta = 0$ ).*

*Then, for all rate constant sets  $\mathcal{K}$  and initial conditions  $\mathbf{x}_0 \in \mathbb{R}_{>0}^m$ , the corresponding mass action system  $(\mathcal{S}, \mathcal{C}, \mathcal{R}, \mathcal{K})$  has the property that there exists a unique positive steady state  $\mathbf{x}^* \in \mathbf{C}_{\mathbf{x}_0}$  and this state is locally asymptotically stable with respect to  $\mathbf{C}_{\mathbf{x}_0}$ .*

This result is surprising since, as just noted, the deficiency depends solely upon the network structure and not upon the choice of kinetics. The result, however, gives conclusions on the admissible dynamics to the corresponding mass action system (1); in fact, it gives very strong conclusions. It is also notable that the result holds independent of the rate constants and initial conditions; that is, it is robust to all of the system's parameter values.

In practice, many reaction networks arising from industrial chemistry and systems biology do not satisfy the assumptions of Theorem 2.1. The following result applies to many networks with a higher deficiency [9, 11].

**Theorem 2.2** (Deficiency One Theorem). *Consider a chemical reaction network  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  with linkage classes  $\mathcal{L} = \{L_1, \dots, L_\ell\}$ . Let  $\delta_\theta$  denote the deficiency of the subnetwork consisting only of the complexes and reactions in  $L_\theta$  for  $\theta = 1, \dots, \ell$ . Suppose that:*

1.  *$\delta_\theta \leq 1$ , for all  $\theta = 1, \dots, \ell$ ,*
2.  *$\sum_{\theta=1}^{\ell} \delta_\theta = \delta$*
3. *Every linkage class contains exactly one terminal strong linkage class (i.e.  $t = \ell$ ).*

*Then, if the mass action system  $(\mathcal{S}, \mathcal{C}, \mathcal{R}, \mathcal{K})$  admits a strictly positive steady state, every stoichiometric compatibility class contains exactly one steady state. Furthermore, if the network is weakly reversible, then  $(\mathcal{S}, \mathcal{C}, \mathcal{R}, \mathcal{K})$  admits a positive steady state for all choices of  $\mathcal{K}$ .*

The conclusions of the Deficiency One Theorem are not as strong as the Deficiency Zero Theorem since it does not give any information about the stability of steady states; in fact, the steady states may be stable or unstable. It should be noted, however, that it is often very difficult to ascertain directly the uniqueness of steady states in compatibility classes and that the Deficiency One Theorem presents a parameter-free method for obtaining this property.

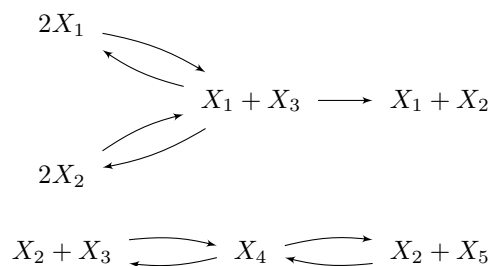
It is also worth noting that, despite the implication of the name, mechanisms satisfying the Deficiency One Theorem *are not* required to have a deficiency of one. For example, it is permissible to have a network with a deficiency of two and two linkage classes with subnetwork deficiencies of one (i.e.  $\delta_1 = 1$ ,  $\delta_2 = 1$ , and  $\delta = \delta_1 + \delta_2 = 2$ ).

Further consequences of condition 2. of the Deficiency One Theorem were considered in [5]. The following result was shown.

**Theorem 2.3** (Corollary 3.5, [5]). *Consider a chemical reaction network  $(S, \mathcal{C}, \mathcal{R})$  for which condition 2. of the Deficiency One Theorem holds. Then, if any stoichiometric compatibility class has a finite number of steady states, every stoichiometric compatibility class has the same finite number of steady states.*

This result eliminates the possibility of bifurcations in the initial conditions resulting in different numbers of steady states. The result again depends only upon the structural information of the network, and not on the parameter values. It is often very challenging to ascertain this information through direct analysis of the differential equations (1).

**Example 2.1.** *Consider the following chemical reaction network:*



*It can be quickly computed that the stoichiometric space has dimension  $s = \dim(S) = 4$  so that the deficiency is  $\delta = n - \ell - s = 7 - 2 - 4 = 1$ . The deficiencies of the two linkage classes, enumerated in the order they appear above, are given by*

$$\begin{aligned}
 \delta_1 &= n_1 - 1 - s_1 = (4) - 1 - (2) = 1 \\
 \delta_2 &= n_2 - 1 - s_2 = (3) - 1 - (2) = 0.
 \end{aligned}$$

*It follows that  $\delta_1 \leq 1$ ,  $\delta_2 \leq 1$ , and  $\delta = \delta_1 + \delta_2 = 1$ . The terminal strong linkage classes are  $\{X_1 + X_2\}$  and  $\{X_2 + X_3, X_4, X_2 + X_5\}$  so that every linkage class only has a single terminal strong linkage class. The Deficiency One Theorem therefore applies so that, if the mass action system  $(S, \mathcal{C}, \mathcal{R}, \mathcal{K})$  admits a positive steady state for some rate constant set  $\mathcal{K}$ , then every stoichiometric compatibility class has exactly one positive steady state.*

### 3 Main Results

In this section, we consider the question of whether, given a mass action system  $(\mathcal{S}, \mathcal{C}, \mathcal{R}, \mathcal{K})$ , we can find a linearly conjugate system  $(\mathcal{S}^*, \mathcal{C}^*, \mathcal{R}^*, \mathcal{K}^*)$  which satisfies the assumptions of the Deficiency One Theorem.

We will show that the answer is a definite yes. We present a MILP framework which checks whether conditions 1., 2. and 3. of the Deficiency One Theorem can be satisfied for a linearly conjugate system. The framework is an extension of recent work by the author and others on various problems within CRNT. The most directly applicable background paper is [22], where the authors present a framework for determining whether a system is linearly conjugate to a system with an underlying network which satisfies the Deficiency Zero Theorem. We also take elements from the recent computational paper [20] which introduces a method for checking conditions on the connectivity of non-weakly reversible networks by relating them to a weakly reversible network.

We will at various points require the following background results from CRNT.

**Theorem 3.1** (Theorem 3.1 of [13]; Proposition 4.1 of [7]). *Let  $A(\mathcal{K})$  be the Laplacian matrix of a mass action system  $(\mathcal{S}, \mathcal{C}, \mathcal{R}, \mathcal{K})$  and let  $\Lambda_\theta$ ,  $\theta = 1, \dots, \ell$ , denote the support of the  $\theta^{\text{th}}$  linkage class,  $L_\theta$ . Then the reaction graph corresponding to  $A(\mathcal{K})$  is weakly reversible if and only if there is a basis of  $\ker(A(\mathcal{K}))$ ,  $\{\mathbf{w}^{(1)}, \dots, \mathbf{w}^{(\ell)}\}$ , such that, for  $\theta = 1, \dots, \ell$ ,*

$$\mathbf{w}^{(\theta)} = \begin{cases} w_j^{(\theta)} > 0, & j \in \Lambda_\theta \\ w_j^{(\theta)} = 0, & j \notin \Lambda_\theta. \end{cases}$$

**Theorem 3.2** (Lemma 4.1, [8]). *Let  $\mathcal{L} = \{L_1, \dots, L_\ell\}$  denote the linkage classes of a chemical reaction network  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ . Then*

$$S = \bigcup_{\theta=1}^{\ell} \text{span}\{y_j - y_i \mid C_i, C_j \in L_\theta\}.$$

**Definition 3.1.** *The **kinetic subspace** of a mass action system  $(\mathcal{S}, \mathcal{C}, \mathcal{R}, \mathcal{K})$  is the smallest subspace of  $\mathbb{R}^m$  which contains  $\text{im}(Y \cdot A(\mathcal{K}) \cdot \Psi(\mathbf{x}(t)))$ .*

**Theorem 3.3** (Unnumbered Theorem, [12]). *Consider a chemical reaction network  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  for which every linkage class contains exactly one terminal strong linkage class (i.e.  $t = \ell$ ). Then, regardless of the choice of  $\mathcal{K}$ , the dimension of the kinetic subspace of the mass action system  $(\mathcal{S}, \mathcal{C}, \mathcal{R}, \mathcal{K})$  is equal to  $s = \dim(S)$ .*

**Lemma 3.1.** *Let  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  denote a chemical reaction network. Suppose  $C'$  is a complex not contained in  $\mathcal{C}$  and  $R'$  is the self-loop  $C' \rightarrow C'$ . Then  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  satisfies the Deficiency One Theorem if and only if  $(\mathcal{S}, \mathcal{C} \cup C', \mathcal{R} \cup R')$  satisfies the Deficiency One Theorem.*

*Proof.* Suppose  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  satisfies the Deficiency One Theorem. The network  $(\mathcal{S}, \mathcal{C} \cup C', \mathcal{R} \cup R')$  differs from  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  only in the isolated complex  $C'$ , which is its own linkage class, and the self-reaction  $R'$ , which does not affect the stoichiometric subspace. It follows that  $n' = n + 1$ ,  $\ell' = \ell + 1$ , and  $s' = s$  where primes are used to denote the network quantities associated with  $(\mathcal{S}, \mathcal{C} \cup C', \mathcal{R} \cup R')$ . We give the isolated complex linkage class the index  $\ell' = \ell + 1$ . It follows that the linkage classes deficiencies coincide for  $\theta = 1, \dots, \ell$ , and that the isolate linkage class gives  $\delta_{\ell'} = n_{\ell'} - 1 - s_{\ell'} =$



$(1) - 1 - (0) = 0 \leq 1$ . We furthermore have that  $\delta = n - \ell - s = (n+1) - (\ell+1) - s = n' - \ell' - s' = \delta'$  so that  $\delta = \sum_{\theta=1}^{\ell} \delta_{\theta} = \sum_{\theta=1}^{\ell'} \delta'_{\theta} = \delta'$  where  $\delta_{\theta} = \delta'_{\theta}$  for  $\theta = 1, \dots, \ell$ , and  $\delta'_{\ell'} = \delta'_{\ell'+1} = 0$ . The isolated complex linkage class clearly also contains a single terminal strong linkage class, corresponding to the complex itself. It follows that  $(\mathcal{S}, \mathcal{C} \cup \mathcal{C}', \mathcal{R} \cup \mathcal{R}')$  satisfies the Deficiency One Theorem. The argument holds in reverse, so that the result is slow.  $\square$

Notice that Lemma 3.1 may be extended to networks with an arbitrary number of isolated complexes added to or removed from a given network.

### 3.1 Mixed-Integer Linear Programming Framework

A MILP problem may be written in the form

$$\begin{aligned} & \text{minimize } \mathbf{c} \cdot \mathbf{x} \\ & \text{subject to } \begin{cases} A_1 \cdot \mathbf{x} = \mathbf{b}_1 \\ A_2 \cdot \mathbf{x} \leq \mathbf{b}_2 \\ x_i \text{ is an integer for } i \in I, I \subseteq \{1, \dots, n\} \end{cases} \end{aligned} \quad (5)$$

where  $\mathbf{x} \in \mathbb{R}^n$  is a vector of unknown decision variables and  $\mathbf{c} \in \mathbb{R}^n$ ,  $\mathbf{b}_1 \in \mathbb{R}^{p_1}$ ,  $\mathbf{b}_2 \in \mathbb{R}^{p_2}$ ,  $A_1 \in \mathbb{R}^{p_1 \times n}$ , and  $A_2 \in \mathbb{R}^{p_2 \times n}$  are vectors and matrices of known parameters values [30].

If all of the decision variables in the problem are real-valued then the problem (5) can be solved in polynomial time. If any of the variables are required to be integer-valued, however, (5) becomes NP-hard. The development of algorithms for efficiently solving MILP problems is a major area of current work which we do not summarize here. For the work in this paper, we utilize the non-commercial software packages GNU Linear Program Kit (GLPK) [26] and SCIP [1].

### 3.2 Initialization of Program

We now set up the objectives of our MILP problem.

Consider two mass action systems  $(\mathcal{S}, \mathcal{C}, \mathcal{R}, \mathcal{K})$  and  $(\mathcal{S}^*, \mathcal{C}^*, \mathcal{R}^*, \mathcal{K}^*)$ . We will refer to the former as the *original system* and the latter as the *target system*. We wish to determine a network structure and associated rate constants for the target network consistent with the requirements of linear conjugacy and either the Deficiency One Theorem or Theorem 2.3. We will suppose that the two networks have common species and complex sets (i.e.  $\mathcal{S} = \mathcal{S}^*$  and  $\mathcal{C} = \mathcal{C}^*$ ), that  $\mathcal{R}$  and  $\mathcal{K}$  are a priori known, and that  $\mathcal{R}^*$  and  $\mathcal{K}^*$  are to be determined. Notice that the first condition implies that the two systems have a common matrix  $Y$  and vector  $\Psi(\mathbf{x})$ . We will assume a priori that the following quantities are known:

- A complex matrix  $Y \in \mathbb{R}_{\geq 0}^{m \times n}$  which is common to both the original and target systems.
- Either the reaction structure  $\mathcal{R}$  and rate constant set  $\mathcal{K}$  of the original system or the numerical values of  $M = Y \cdot A(\mathcal{K})$  corresponding to the coefficient map of  $\Psi(\mathbf{x})$  in the mass action system (1). Notice that  $M$  may be determined from  $Y$  and  $\mathcal{K}$  if they are known.
- The dimension of the kinetic space of the original system, which we will denote by  $s$ . This value is easily computed as the rank of the matrix  $M$ . The equivalence of the kinetic subspace of the original network and stoichiometric subspace of the target system follows from condition 3. of the Deficiency One Theorem and Theorem 3.3.

- A small parameter  $\epsilon > 0$ .
- A set of random variables  $\delta[i, j]$  chosen uniformly from the range  $[\sqrt{\epsilon}, 1/\sqrt{\epsilon}]$ .

Since we do not know how many linkage classes the target system will contain, but must know this to calculate the deficiency by Definition 2.3, we will use the upper bound  $n - s$  (see [22]). We will also allow the networks to contain unused complexes and note that, if a complex does not appear in the target system, it may be treated as a single isolated linkage class with a self-reaction. By Lemma 3.1, this does not affect our ability to apply the Deficiency One Theorem.

### 3.3 Implementing Dynamical Equivalence and Linear Conjugacy

In this section, we introduce constraint sets which guarantee that the original and target systems are linearly conjugate to one another. A full description of the process can be found in [21] and [24].

We introduce the following decision variables:

$$\begin{cases} d[i] > \epsilon, & i = 1, \dots, m \\ b[i, j] \in [0, \frac{1}{\epsilon}], & i, j = 1, \dots, n, i \neq j. \end{cases} \quad (\text{Dec1})$$

where the  $d[i]$  correspond to the reciprocals of the conjugacy constants  $c_i$  (i.e.  $d[i] = 1/c_i$ ) and the  $b[i, j]$  correspond to scalings of the rate constants  $k^*(i, j)$  of the target network.

To see why we track  $d[i]$  and  $b[i, j]$ , rather than  $c_i$  and  $k^*(i, j)$ , respectively, we briefly reproduce the arguments of [21]. We start by defining  $T = \text{diag}\{\mathbf{c}\}$ . We can write the conjugacy transformation as  $\mathbf{x} = T \cdot \mathbf{x}^*$  so that the two networks are linearly conjugate for rate constant set  $\mathcal{K}^*$  if

$$Y \cdot A(\mathcal{K}) \cdot \Psi(\mathbf{x}) = T \cdot Y \cdot A(\mathcal{K}^*) \cdot \Psi(\mathbf{x}^*) = T \cdot A(\mathcal{K}^*) \cdot \text{diag}\{\Psi(\mathbf{c})\} \cdot \Psi(\mathbf{x}).$$

We can simplify this expression by removing  $\Psi(\mathbf{x})$  and making the substitutions  $A(\mathcal{B}) = A(\mathcal{K}^*) \cdot \text{diag}\{\Psi(\mathbf{c})\}$  and  $M = Y \cdot A(\mathcal{K})$  (see Theorem 2 of [24]). After inverting  $T$ , we can see that the two networks are linearly conjugate if they satisfy the following linear constraint set:

$$\begin{cases} Y \cdot A(\mathcal{B}) = T^{-1} \cdot M. \end{cases} \quad (\text{LC})$$

$Y$  and  $M$  are a priori known while  $A(\mathcal{B})$  and  $T^{-1}$  contain the unknown decision variables  $b[i, j]$  and  $d[i]$  (since  $T^{-1} = \text{diag}\{1/\mathbf{c}\}$ ). Notice that the conjugacy constants can be obtained from  $c_i = 1/d[i]$  and the rate constant set  $\mathcal{K}^*$  can then be recovered by the equation

$$A(\mathcal{K}^*) = A(\mathcal{B}) \cdot [\text{diag}\{\Psi(\mathbf{c})\}]^{-1}.$$

The distribution of positive and zero elements for  $b[i, j]$  and  $k^*[i, j]$  coincide so that the two networks correspond to the same network structure  $\mathcal{R}^*$ .

### 3.4 Implementing Deficiency Conditions

In this section, we incorporate conditions 1. and 2. of the Deficiency One Theorem into a mixed integer linear programming framework. To accommodate Definition 2.3, we need to be able to

track the linkage classes and stoichiometric subspaces of both the original and target systems. We introduce the following decision variables:

$$\left\{ \begin{array}{ll} \Lambda[i, \theta] \in \{0, 1\}, & i = 1, \dots, n, \theta = 1, \dots, n-s \\ \Gamma[i, j, \theta] \in \{0, 1\}, & i, j = 1, \dots, n, i \neq j, \theta = 1, \dots, n-s \\ S[i, j, \theta] \geq 0, & i, j = 1, \dots, n, i \neq j, \theta = 1, \dots, n-s \\ S'[i, j, \theta] \in \{0, 1\}, & i, j = 1, \dots, n, i \neq j, \theta = 1, \dots, n-s \\ L[\theta] \in [0, 1], & \theta = 1, \dots, n-s. \end{array} \right. \quad (\text{Dec2})$$

*Counting linkage classes:* We will count the linkage classes of the target network using the techniques presented in [22]. To track the linkage classes of the target system, we desire the following logical equivalences:

$$\begin{aligned} \Lambda[i, \theta] = 1 &\iff C_i \in L_\theta^* \\ L[\theta] = 1 &\iff L_\theta^* \neq \emptyset \end{aligned}$$

where  $L_\theta^*, \theta = 1, \dots, n-s$ , are the linkage classes of the target system. We can accomplish this with the following set of constraints:

$$\left\{ \begin{array}{ll} b[i, j] \leq \frac{1}{\epsilon} \cdot (\Lambda[i, \theta] - \Lambda[j, \theta] + 1), & i, j = 1, \dots, n, i \neq j, \theta = 1, \dots, n-s \\ \sum_{\theta=1}^{n-s} \Lambda[i, \theta] = 1, & i = 1, \dots, n \\ \sum_{i=1}^n \Lambda[i, \theta] - \epsilon \cdot L[\theta] \geq 0, & \theta = 1, \dots, n-s \\ -\sum_{i=1}^n \Lambda[i, \theta] + \frac{1}{\epsilon} \cdot L[\theta] \geq 0, & \theta = 1, \dots, n-s \\ \sum_{j=1}^i \Lambda[j, \theta] \geq \sum_{l=\theta+1}^{n-s} \Lambda[i, l], & i = 1, \dots, n, \theta = 1, \dots, n-s, \theta \leq i \end{array} \right. \quad (\text{Linkage})$$

The first constraint guarantees that the scaled rate constant  $b[i, j]$  is zero if  $C_i$  and  $C_j$  do not belong to the same linkage class  $L_\theta$  in the target network. The second constraint guarantees that every complex is assigned to exactly one linkage class. The third and fourth constraints guarantee that  $L[\theta]$  is zero if  $L_\theta^* = \emptyset$  and one if  $L_\theta^* \neq \emptyset$ . That is, it counts the number of linkage classes in the network. The fifth constraint removes redundant permutations in the assignment of complexes to linkage classes and is necessary for computational efficiency (see [22] for full justification).

*Computing stoichiometric subspace dimensions:* In order to determine the deficiency of each linkage class we need to compute  $\delta_\theta = n_\theta - 1 - s_\theta$  where  $n_\theta$  is the number of complexes in  $L_\theta^*$  and  $s_\theta$  is the dimension of the stoichiometric subspace of the corresponding subnetwork.

We will compute  $s_\theta$  by using the a priori known random variables  $\delta[i, j]$  to construct a random vector in the span of the reaction vectors on the support on  $L_\theta^*$ . Notice that this spans  $S_\theta$  by Theorem 3.2. With probability one, the dimension will correspond to the minimal number of vectors required to reach this random vector through a linear combination. We want the following

logical equivalences:

$$\begin{aligned}\Gamma[i, j, \theta] = 1 &\iff C_i, C_j \in L_\theta^* \\ S'[i, j, \theta] = 1 &\iff y_j - y_i \text{ is a basis element of } S_\theta \\ S[i, j, \theta] > 0 &\iff y_j - y_i \text{ is a basis element of } S_\theta.\end{aligned}$$

We can accomplish this with the following constraints:

$$\left\{ \begin{array}{ll} S'[i, j, \theta] \leq \Gamma[i, j, \theta], & i, j = 1, \dots, n, i \neq j, \theta = 1, \dots, n-s \\ S[i, j, \theta] \leq \frac{1}{\epsilon} \cdot S'[i, j, \theta], & i, j = 1, \dots, n, i \neq j, \theta = 1, \dots, n-s \\ -S[i, j, \theta] \leq -\epsilon \cdot S'[i, j, \theta], & i, j = 1, \dots, n, i \neq j, \theta = 1, \dots, n-s \\ \Gamma[i, j, \theta] \leq 1 + \epsilon \cdot (\Lambda[i, \theta] + \Lambda[j, \theta] - 2), & i, j = 1, \dots, n, i \neq j, \theta = 1, \dots, n-s \\ \Gamma[i, j, \theta] \geq \epsilon \cdot (\Lambda[i, \theta] + \Lambda[j, \theta] - 1), & i, j = 1, \dots, n, i \neq j, \theta = 1, \dots, n-s \\ \sum_{\substack{i,j=1 \\ i \neq j}}^n S[i, j, \theta] \cdot (Y[k, j] - Y[k, i]) \\ = \sum_{\substack{i,j=1 \\ i \neq j}}^n \Gamma[i, j, \theta] \cdot \delta[i, j] \cdot (Y[k, j] - Y[k, i]), & \theta = 1, \dots, n-s, k = 1, \dots, n \end{array} \right. \quad (\text{Stoic})$$

The first constraint restricts the basis vectors to those on the same linkage class. The second and third constraints guarantee that  $S[i, j, \theta] \in [\epsilon, 1/\epsilon]$  if  $S'[i, j, \theta] = 0$  and  $S[i, j, \theta] = 0$  if  $S'[i, j, \theta] = 0$ . The fourth and fifth constraints guarantee that  $\Gamma[i, j, \theta] = 1$  if and only if  $C_i \in L_\theta^*$  and  $C_j \in L_\theta^*$ . The final constraint determines, for each linkage class, the number of vectors requires to reach a random vector in the corresponding subspace.

*Conditions 1. and 2. of the Deficiency One Theorem:* We can now accommodate conditions 1. and 2. of the Deficiency One Theorem with the following constraint sets:

$$\left\{ \begin{array}{l} -\sum_{\substack{i,j=1 \\ i \neq j}}^n S'[i, j, \theta] \leq 2 - \sum_{i=1}^n \Lambda[i, \theta], \quad \theta = 1, \dots, n-s \\ \sum_{\theta=1}^{n-s} \sum_{\substack{i,j=1 \\ i \neq j}}^n S'[i, j, \theta] = s. \end{array} \right. \quad (\text{DOT})$$

The first constraint guarantees that  $\delta_\theta \leq 1$  for all  $\theta = 1, \dots, n-s$ , while the second constraint guarantees that  $\sum_{\theta=1}^{n-s} s_\theta = s$  so that  $\sum_{\theta=1}^{n-s} \delta_\theta = \sum_{\theta=1}^{n-s} (n_\theta - 1 - s_\theta) = n - \ell - s = \delta$ . We can accommodate the simpler conditions of Theorem 2.3 by using the following constraint set as an alternative to (DOT):

$$\left\{ \begin{array}{l} \sum_{\theta=1}^{n-s} \sum_{\substack{i,j=1 \\ i \neq j}}^n S'[i, j, \theta] = s. \end{array} \right. \quad (\text{Boros})$$

### 3.5 Implementing One Terminal Strong Linkage Class

In order to restrict the target network to contain only a single terminal strong linkage class in each linkage class, we introduce the following result.

**Lemma 3.2.** *Consider a chemical reaction network  $(S, \mathcal{C}, \mathcal{R})$  with linkage classes  $\mathcal{L} = \{L_1, \dots, L_\ell\}$ . Then every linkage class contains only a single terminal strong linkage class if and only if there is a set of complexes  $\mathcal{C}' \subseteq \mathcal{C}$  and supplemental set of reactions  $\mathcal{R}' \subseteq \mathcal{C} \times \mathcal{C}$  such that:*

1.  $|\mathcal{C}' \cap L_\theta| \leq 1$  for every  $\theta = 1, \dots, \ell$ ;
2.  $(C_i, C_j) \in \mathcal{R}'$  implies that  $C_i \in \mathcal{C}'$  and  $C_i, C_j \in L_\theta$  for some  $\theta \in \{1, \dots, \ell\}$ .
3. The network  $(S, \mathcal{C}, \mathcal{R} \cup \mathcal{R}')$  is weakly reversible.

*Proof.* ( $\implies$ ) Suppose that the network has a single terminal strong linkage class in each linkage class. Let  $\mathcal{C}'$  contain a single complex from each terminal strong linkage class and  $\mathcal{R}'$  consist of all reactions from a complex in  $\mathcal{C}'$  to a complex in the same linkage class. Then conditions 1., 2., and 3. are trivially satisfied.

( $\impliedby$ ) Suppose that at least one linkage class has multiple terminal strong linkage classes. It follows by condition 1. and 2. that we may only introduce reactions which lead from a single terminal strong linkage class to any other complex. The remaining terminal strong linkage classes must remain terminal. It follows that we cannot satisfy condition 3. and we are done.  $\square$

This result says that the condition that every linkage class contains a single terminal strong linkage class can only be satisfied if supplemental reactions from at most one complex in each terminal strong linkage class can make the network weakly reversible. We will use this result in conjunction with Theorem 3.1 to identify networks satisfying the Deficiency One Theorem.

To accommodate the conditions of Lemma 3.2, and therefore condition 3. of the Deficiency One Theorem, we introduce the decision variables:

$$\left\{ \begin{array}{ll} w[i, j] \geq 0, & i, j = 1, \dots, n, \ i \neq j \\ w'[i, j] \geq 0, & i, j = 1, \dots, n, \ i \neq j \\ C[i, \theta] \in \{0, 1\}, & i = 1, \dots, n, \ \theta = 1, \dots, n-s \\ C'[i] \in [0, 1], & i = 1, \dots, n. \end{array} \right. \quad (\text{Dec3})$$

*Condition 1. of Lemma 3.2:* We wish to impose the following logical equivalences:

$$\begin{aligned} C[i, \theta] = 1 &\implies C_i \in \mathcal{C}' \text{ and } C_i \in L_\theta \\ C[i] = 1 &\iff C_i \in \mathcal{C}' \\ |\mathcal{C}' \cap L_\theta| \leq 1 &\text{ for every } \theta = 1, \dots, \ell. \end{aligned}$$

We can guarantee this with the following constraint set:

$$\left\{ \begin{array}{ll} C[i, \theta] \leq \Lambda[i, \theta], & i = 1, \dots, n, \ \theta = 1, \dots, n-s \\ C'[i] = \sum_{\theta=1}^{n-s} C[i, \theta], & i = 1, \dots, n \\ \sum_{i=1}^n C[i, \theta] \leq 1, & \theta = 1, \dots, n-s. \end{array} \right. \quad (\text{C'})$$

The first condition guarantees that the intermediate decision variable  $C[i, \theta]$  attains the value one only if  $C_i \in L_\theta$ . The second condition determines whether  $C_i \in \mathcal{C}'$ . The third condition guarantees that each linkage class of the target network contains at most one such complex.

*Conditions 2. and 3. of Lemma 3.2:* We wish to impose the following:

$$\begin{aligned} w[i, j] > 0 &\iff b[i, j] > 0 \\ w'[i, j] > 0 &\iff (C_i, C_j) \in \mathcal{R}' \\ (\mathcal{S}^*, \mathcal{C}^*, \mathcal{R}^* \cup \mathcal{R}') &\text{ is weakly reversible.} \end{aligned}$$

This can be accomplished with the following constraint sets:

$$\left\{ \begin{array}{ll} w[i, j] - \epsilon \cdot b[i, j] \geq 0, & i, j = 1, \dots, n, i \neq j \\ -w[i, j] + \frac{1}{\epsilon} \cdot b[i, j] \geq 0, & i, j = 1, \dots, n, i \neq j \\ w'[i, j] \leq C'[i], & i, j = 1, \dots, n, i \neq j. \\ w'[i, j] \leq \frac{1}{\epsilon} \cdot (\Lambda[i, \theta] - \Lambda[j, \theta] + 1), & i, j = 1, \dots, n, i \neq j, \theta = 1, \dots, n - s \\ \sum_{\substack{j=1 \\ j \neq i}}^n (w[i, j] + w'[i, j]) = \sum_{\substack{j=1 \\ j \neq i}}^n (w[j, i] + w'[j, i]), & i = 1, \dots, n \end{array} \right. \quad (\text{WR})$$

The first two constraints guarantee that  $w[i, j]$  and  $b[i, j]$  have the same distribution of positive and zero elements. The third constraint guarantees only reactions from complexes in  $\mathcal{C}'$  are allowed to be included in  $\mathcal{R}'$ , while the fourth constraint guarantees only reactions within linkage classes may be included in  $\mathcal{R}'$ . The final constraint guarantees the network  $(\mathcal{S}^*, \mathcal{C}^*, \mathcal{R}^* \cup \mathcal{R}')$  is weakly reversible according to Theorem 3.1 (see [24]).

*Objective function:* The existence of a target network satisfying the requirements of the Deficiency One Theorem is only dependent upon the feasible region of the MILP being non-empty. We still need, however, an objective function. It is often convenient to maximize the number of linkage classes, which minimizes the deficiency of the target network (see [22]). This can be accomplished with:

$$\text{minimize } \left\{ - \sum_{\theta=1}^{n-s} L[\theta] \right. \quad (\text{Obj})$$

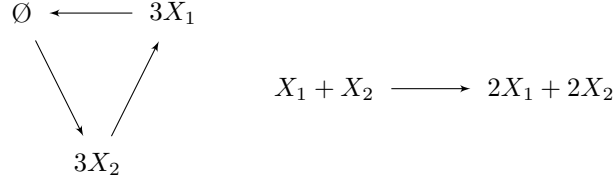
A target network satisfying the requirements of the Deficiency One Theorem can be found by optimizing **(Obj)** over the decision variables **(Dec1)**, **(Dec2)**, and **(Dec3)**, and the constraint sets **(LC)**, **(Linkage)**, **(Stoic)**, **(DOT)**, **(C')**, and **(WR)**. Theorem 2.3 may be checked instead of the Deficiency One Theorem by replacing **(DOT)** with **(Boros)**.

## 4 Examples

In this example, we apply the algorithm outlined in Section 3 to example mass action systems. All computations were performed on the author's professional use HP Spectre 360 laptop (Intel Core i7-5500U CPU @ 2.40 GHz, 8.00 GB RAM). The optimization programs used were GLPK [26] and

SCIP [1].

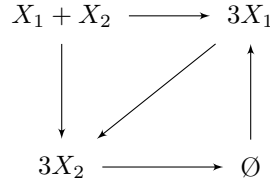
**Example 4.1.** Consider the following chemical reaction network:



where  $\emptyset$  corresponds to the zero complex, which has all zero stoichiometric coefficients. The zero complex is commonly used in CRNT to denote inflows and outflows of species in the system.

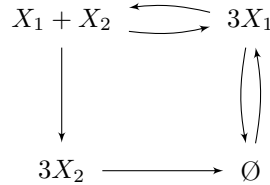
This network is not directly amenable to the Deficiency One Theorem because we have  $\delta = 1$  for the entire network but  $\delta_1 = \delta_2 = 0$  for the two linkage class subnetworks so that  $\delta \neq \delta_1 + \delta_2$ . We will run the algorithm contained in Section 3 twice, both times taking all rate constants  $k(i, j) = 1$  and  $\epsilon = 0.1$ .

*Dynamical equivalency (i.e.  $c_i = 1$  for  $i = 1, 2$ ):* The algorithm produces the following network:



with all rate constants  $k^*(i, j) = 1$ . We have excluded the isolated complex  $2X_1 + 2X_2$  and associated self-reaction since it is not relevant to our study by Lemma 3.1. It can be easily checked that shown linkage classes has a single terminal strong linkage class consisting of  $\{\emptyset, 3X_1, 3X_2\}$ , and that the deficiency is  $\delta = 1$ . Since the network consists of only the single linkage class, it follows that the system satisfies the Deficiency One Theorem.

*Non-trivial linear conjugacy (i.e.  $c_i$  varying):* The algorithm produces the following network:



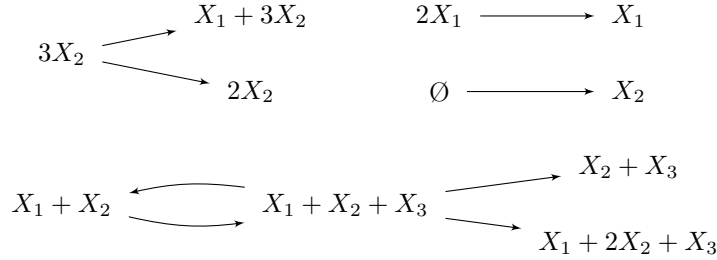
with the values  $c_1 = 1$ ,  $c_2 = 4.66666$ . To state the rate constants, we set  $C_1 = \emptyset$ ,  $C_2 = 3X_1$ ,  $C_3 = 3X_2$ , and  $C_4 = X_1 + X_2$ . The program gives the rate constants  $k^*(1, 2) = 1$ ,  $k^*(2, 1) = 0.571429$ ,  $k^*(2, 4) = 0.642857$ ,  $k^*(3, 1) = 0.002109$ ,  $k^*(4, 2) = 0.158163$ ,  $k^*(4, 3) = 0.102041$ , and the rest  $k^*(i, j) = 0$ .

We again have that the network satisfies the Deficiency One Theorem. It is worth noting that this network is weakly reversible while insisting on dynamical equivalence yields a network which was not weakly reversible. This weak reversibility allows us to conclude that there does, in fact, exist a positive steady state, and so this steady state must be unique in the whole state space  $\mathbb{R}_{>0}^2$  since  $S = \mathbb{R}^2$ .

**Example 4.2.** Consider the following mass action system:

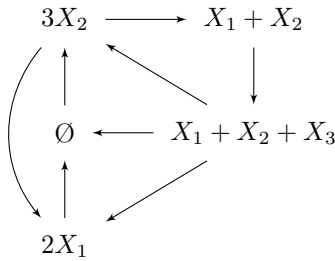
$$\begin{aligned}\dot{x}_1 &= 2x_2^3 - x_1^2 - x_1x_2x_3 \\ \dot{x}_2 &= 1 - 3x_2^3 + 3x_1x_2x_3 \\ \dot{x}_3 &= x_1x_2 - x_1x_2x_3.\end{aligned}\tag{6}$$

This system can be converted into a chemical reaction network by the algorithm presented in [15] and adapted to optimizing reaction network structures in [32]. We will not reproduce the algorithm here. It yields the following chemical reaction network:



This network is not weakly reversible, and has two linkage classes which have multiple terminal strong linkage classes. The overall deficiency is  $\delta = 4$  while the deficiency of each linkage class is  $\delta_\theta = 0$  for  $\theta = 1, \dots, 4$ . It follows that neither the Deficiency One Theorem nor Theorem 2.3 apply to this network.

We now apply the algorithm provided in Section 3 to see if there is a linearly conjugate system satisfying either of these theorems. We take  $\epsilon = 0.1$ . This gives the following network:



where  $c_1 = 1$ ,  $c_2 = 0.9941642558$ , and  $c_3 = 0.2840295845$ , and the remainder of the complexes are only involved in self-reactions (omitted). We index the complexes as  $C_1 = \emptyset$ ,  $C_2 = 3X_2$ ,  $C_3 = 2X_1$ ,  $C_4 = X_1 + X_2$ , and  $C_5 = X_1 + X_2 + X_3$ . The algorithm gives the rate constants  $k^*(1, 2) = 0.335289$ ,  $k^*(2, 3) = 0.999798$ ,  $k^*(2, 4) = 0.358313$ ,  $k^*(3, 1) = 0.5$ ,  $k^*(4, 5) = 3.541427$ ,  $k^*(5, 1) = 0.286590$ ,  $k^*(5, 2) = 7.718363$ , and  $k^*(5, 3) = 4.463544$ .



*This network is weakly reversible and has a single linkage class. The deficiency of the network is  $\delta = 1$  so that the Deficiency One Theorem applies. Since the stoichiometric subspace is  $S = \mathbb{R}_{>0}^3$ , it follows that there is a unique positive steady state for the system (6) in  $\mathbb{R}_{>0}^3$ . It is worth noting here that restricting to dynamical equivalence, and not full linear conjugacy, does not produce a network satisfying either the Deficiency Zero Theorem or Theorem 2.3.*

## 5 Conclusions

In this paper, we have presented a MILP framework for determining whether a given mass action system is linearly conjugate to a system which satisfies the well-known Deficiency One Theorem or the recent generalization stated here as Theorem 2.3. In particular, we have outlined constraint sets capable of imposing the following critical assumption: (a) that the sum of the deficiencies of each linkage classes is bounded by one and sums to the overall network deficiency; and (b) that each linkage class contains a single terminal strong linkage class. We also presented examples of systems for which the Deficiency One Theorem could be applied only after the algorithm was utilized.

This paper raises some interesting avenues for future work:

1. *Determination of optimal complex set:* Current MILP algorithms for determining optimal network structures within CRNT require that the complex set  $\mathcal{C}$ , and therefore  $Y$ , be specified prior to application of the algorithm. An initial complex set which is too small, however, runs the risk of not finding an admissible network, while a complex set which is too large increases computationally inefficiency. The network determination algorithm presented in [15], for instance, typically produces complex sets which are unnecessarily large (e.g. only 5 of the 11 complexes in Example 4.2 were needed in the linearly conjugate system). Determining methods of complex selection which bound the size of  $Y$  is therefore a primary concern for future research.
2. *Parameter-free approach:* The MILP algorithm outlined in this paper depends upon the rate constants of the original network being specified. It is often beneficial, however, to leave this set unspecified; that is, to search over all possible mass action systems associated with a given network structure for a linearly conjugate system satisfying the Deficiency One Theorem. A parameter-free approach was introduced in [23] but is only known to be linear for dynamically equivalent relationships and not fully linearly conjugate ones. We saw in Example 4.2, however, that the full application of linearly conjugate may be required to apply the Deficiency One Theorem. Extending the underlying theory to incorporate parameter-free approaches will therefore be the focus of future work.

**Acknowledgments:** The author gratefully thanks San José State University for its financial and logistic support, and Lake Tahoe for the serenity which lead to the conception of this project.

## References

- [1] Tobias Achterberg. SCIP: Solving constraint integer programs. *Math. Program. Comput.*, 1(1):1–41, 2009. <http://mpc.zib.de/index.php/MPC/article/view/4>
- [2] Bernadett Ács, Gábor Szederkényi, Zoltán A. Tuza, and Zsolt Tuza. Computing linearly conjugate weakly reversible kinetic structures using optimization and graph theory. *MATCH Commun. Math. Comput. Chem.*, 74:489–512, 2015.

- [3] Carlene Perpetua P. Arceo, Editha C. Jose, Alberto Marin-Sanguino, and Eduardo R. Mendoza. Chemical reaction network approaches to biochemical systems theory. *Math. Biosci.*, 269:135–152, 2015.
- [4] Balázs Boros. Notes on the deficiency-one theorem: Multiple linkage classes. *Math. Biosci.*, 235(1):110–122, 2012.
- [5] Balázs Boros. On the positive steady states of deficiency-one mass action systems, PhD Thesis, Eotvos Loránd University, Budapest, 2013.
- [6] Gheorghe Craciun and Casian Pantea. Identifiability of chemical reaction networks. *J. Math Chem.*, 44(1):244–259, 2008.
- [7] Martin Feinberg. Lectures on chemical reaction networks. Unpublished written versions of lectures given at the Mathematics Research Center, University of Wisconsin, 1979. <https://crnt.osu.edu/LecturesOnReactionNetworks>
- [8] Martin Feinberg. Complex balancing in general kinetic systems. *Arch. Ration. Mech. Anal.*, 49:187–194, 1972.
- [9] Martin Feinberg. Chemical reaction network structure and the stability of complex isothermal reactors: I. The deficiency zero and deficiency one theorems. *Chem. Eng. Sci.*, 42(10):2229–2268, 1987.
- [10] Martin Feinberg. Chemical reaction network structure and the stability of complex isothermal reactors: II. Multiple steady states for networks of deficiency one. *Chem. Eng. Sci.*, 43(1):1–25, 1988.
- [11] Martin Feinberg. The existence and uniqueness of steady states for a class of chemical reaction networks. *Arch. Ration. Mech. Anal.*, 132:311–370, 1995.
- [12] Martin Feinberg and Fritz Horn. Chemical mechanism structure and the coincidence of the stoichiometric and kinetic subspaces. *Arch. Rational Mech. Anal.*, 66:83–97, 1977.
- [13] Karin Gatermann and Birkett Huber. A family of sparse polynomial systems arising in chemical reaction systems. *J. Symbolic Comput.*, 33(3):275–305, 2002.
- [14] Jeremy Gunawardena. Chemical reaction network theory for in-silico biologists. Unpublished lecture notes, 2003. <http://vcp.med.harvard.edu/papers/crnt.pdf>
- [15] V. Hárs and János Tóth. On the inverse problem of reaction kinetics. *Coll. Math. Soc. J. Bolyai*, 30:363–379, 1981.
- [16] Archibald Hill. The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. *J. Physiol.*, 40(4), 2010.
- [17] Fritz Horn. Necessary and sufficient conditions for complex balancing in chemical kinetics. *Arch. Ration. Mech. Anal.*, 49:172–186, 1972.
- [18] Fritz Horn and Roy Jackson. General mass action kinetics. *Arch. Ration. Mech. Anal.*, 47:81–116, 1972.
- [19] Haixia Ji. *Uniqueness of equilibria for complex chemical reaction networks*. PhD thesis, The Ohio State University, 2011.
- [20] Matthew D. Johnston. A computational approach to steady state correspondence of regular and generalized mass action systems. *Bull. Math. Biol.*, 77(6):1065–1100, 2015.
- [21] Matthew D. Johnston and David Siegel. Linear conjugacy of chemical reaction networks. *J. Math. Chem.*, 49(7):1263–1282, 2011.

- [22] Matthew D. Johnston, David Siegel, and Gábor Szederkényi. Computing weakly reversible linearly conjugate chemical reaction networks with minimal deficiency. *Math. Biosci.*, 50(1):274–288, 2012.
- [23] Matthew D. Johnston, David Siegel, and Gábor Szederkényi. Dynamical equivalence and linear conjugacy of chemical reaction networks: New results and methods. *MATCH Commun. Math. Comput. Chem.*, 68(2):443–468, 2012.
- [24] Matthew D. Johnston, David Siegel, and Gábor Szederkényi. A linear programming approach to weak reversibility and linear conjugacy of chemical reaction networks. *J. Math. Chem.*, 50(1):274–288, 2012.
- [25] Gyorgy Lipták, Gábor Szederkényi, and Katalin M. Hangos. Computing zero deficiency realizations of kinetics systems. *Systems Control Lett.*, 81:24–30, 2015.
- [26] Andrew Makhorin. *GNU Linear Programming Kit Reference Manual Version 4.45*. 2010. <http://kam.mff.cuni.cz/~elias/glpk.pdf>.
- [27] Leonor Michaelis and Maud Menten. Die kinetik der invertinwirkung. *Biochem. Z.*, 49:333–369, 1913.
- [28] János Rudan, Gábor Szederkényi, Katalin M. Hangos, and Tamás Péni. Polynomial time algorithms to determine weakly reversible realizations of chemical reaction networks. *J. Math. Chem.*, 52(5):1386–1404, 2014.
- [29] Guy Shinar and Martin Feinberg. Structural sources of robustness in biochemical reaction networks. *Science*, 327(5971):1389–1391, 2010.
- [30] Gabor Szederkényi. Computing sparse and dense realizations of reaction kinetic systems. *J. Math. Chem.*, 47:551–568, 2010.
- [31] Gabor Szederkényi, Julio R. Banga, and Antonio Alvarez-Alonso. CRNreals: a toolbox for distinguishability and identifiability analysis of biochemical reaction networks. *Bioinformatics*, 28(11):1549–1550, 2012.
- [32] Gabor Szederkényi and Katalin Hangos. Finding complex balanced and detailed balanced realizations of chemical reaction networks. *J. Math. Chem.*, 49:1163–1179, 2011.
- [33] Gabor Szederkényi, Katalin Hangos, and Tamas Péni. Maximal and minimal realizations of chemical kinetics systems: computation and properties. *MATCH Commun. Math. Comput. Chem.*, 65:309–332, 2011.
- [34] Gabor Szederkényi, Katalin Hangos, and Zsolt Tuza. Finding weakly reversible realizations of chemical reaction networks using optimization. *MATCH Commun. Math. Comput. Chem.*, 67:193–212, 2012.
- [35] Aizik I. Vol’pert and Sergei I. Hudjaev. *Analysis in Classes of Discontinuous Functions and Equations of Mathematical Physics*. Martinus Nijhoff Publishers, Dordrecht, Netherlands, 1985.